Supplementary Information (SI) Appendix for
Gut microbiome structure and metabolic activity in inflammatory bowel disease

Eric A. Franzosa[1,2,*], Alexandra Sirota-Madi[1,*], Julian Avila-Pacheco[1], Nadine Fornelos[1], Henry J. Haiser[3], Stefan Reinker[3], Tommi Vatanen[1], A. Brantley Hall[1], Himel Mallick[1,2], Lauren J. McIver[1,2], Jenny S. Sauk[4], Robin G. Wilson[4], Betsy W. Stevens[4], Justin M. Scott[1], Kerry Pierce[1], Amy A. Deik[1], Kevin Bullock[1], Floris Imhann[5,6], Jeffrey Porter[3], Alexandra Zhernakova[6], Jingyuan Fu[6,7], Rinse K. Weersma[5], Cisca Wijmenga[6,8], Clary B. Clish[1], Hera Vlamakis[1], Curtis Huttenhower[1,2,†], Ramnik J. Xavier[1,4,9,†]

[1]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

[2]Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA

[3]Novartis Institute for Biomedical Research Inc., Cambridge, MA 02139, USA

[4]Gastrointestinal Unit and Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

[5]University of Groningen and University Medical Center Groningen, Department of Gastroenterology and Hepatology, Groningen, the Netherlands

[6]University of Groningen and University Medical Center Groningen, Department of Genetics, Groningen, the Netherlands

[7]University of Groningen and University Medical Center Groningen, Department of Pediatrics, Groningen, the Netherlands

[8]K.G. Jebsen Coeliac Disease Research Centre, Department of Immunology, University of Oslo, Norway

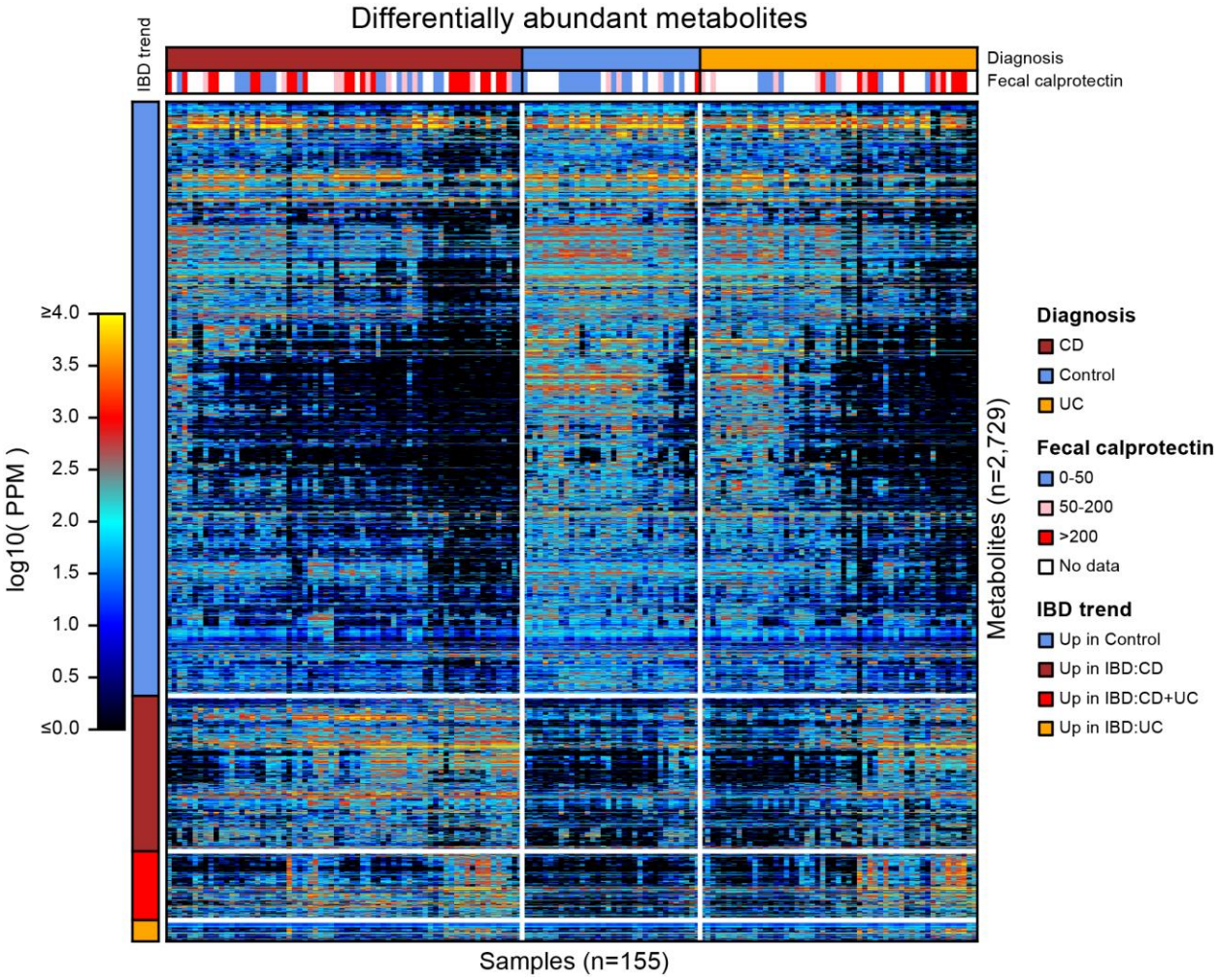
[9]Center for Microbiome Informatics and Therapeutics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*These authors contributed equally to this work

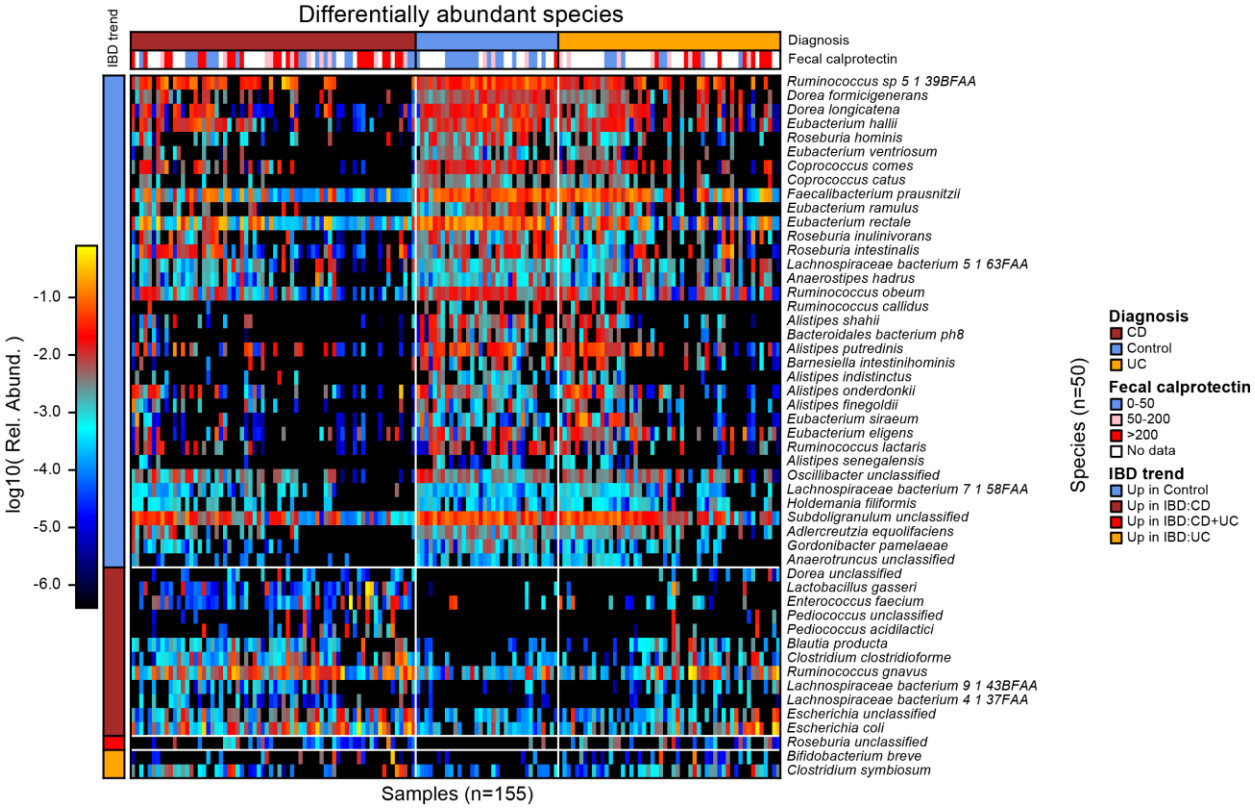
[†]Corresponding authors: xavier@molbio.mgh.harvard.edu; chuttenh@hsph.harvard.edu

Supplementary Figures

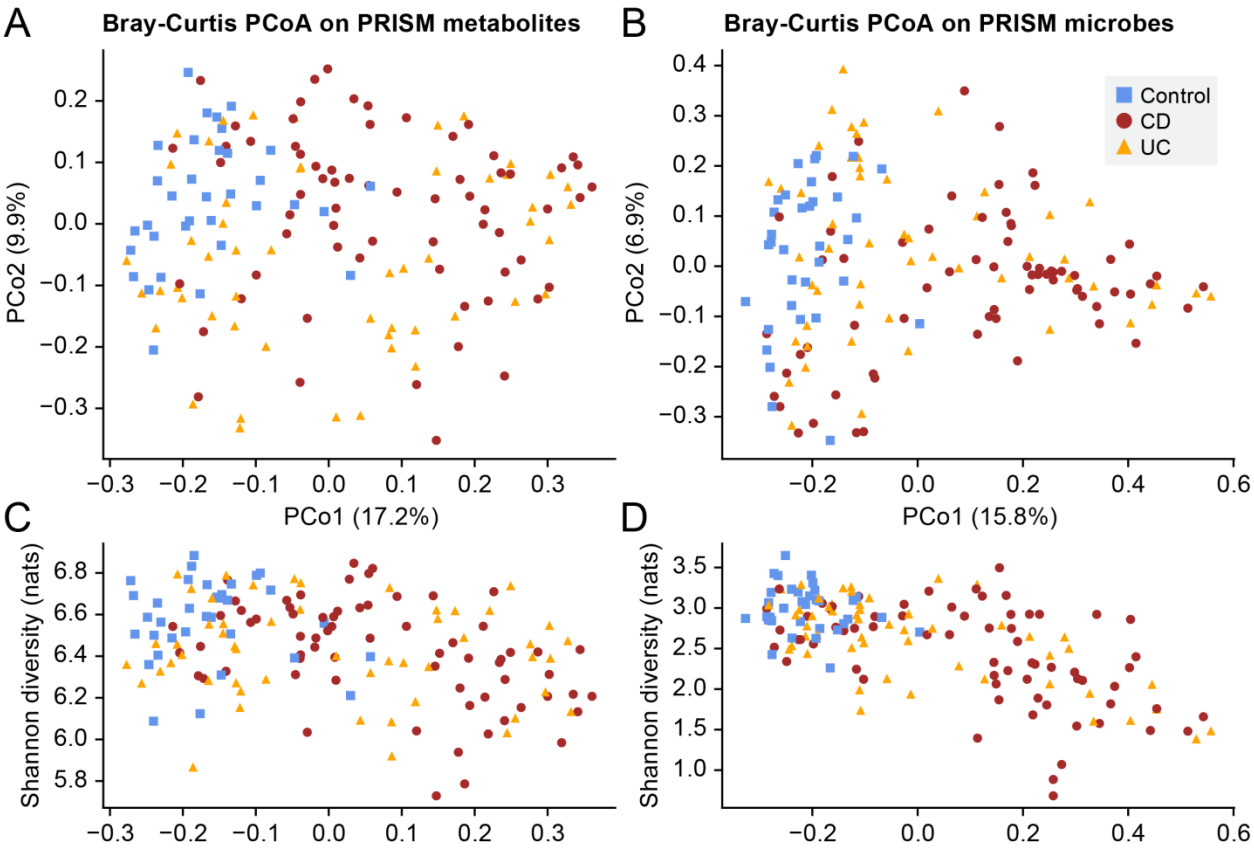
Supplementary Figure 1: Differentially abundant metabolites in IBD. 2,729 metabolite features were differentially abundant in IBD (UC and/or CD) relative to controls as determined by a linear model controlling for subject age and medication use (see main **Methods**). Samples (subjects) were clustered by Bray-Curtis similarity and then grouped by diagnosis. This ordering of subjects is used in all other heatmap figures. A subpopulation of UC subjects (clustered toward the left) more closely resembled non-IBD controls. These subjects also tended to have more control-like levels of inflammation, as measured by fecal calprotectin level. Rows (metabolite features) were clustered by Spearman correlation, then grouped by trend with IBD. Units are “parts per million (PPM)” derived from sum-normalizing metabolomic features intensities within each LC-MS method.



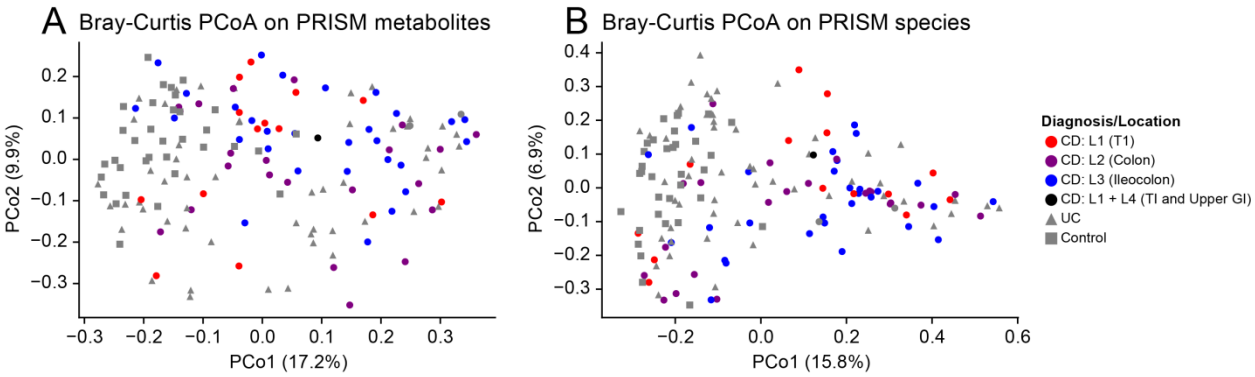
Supplementary Figure 2: Differentially abundant species in IBD. 50 microbial species were differentially abundant in IBD (UC and/or CD) relative to controls as determined by a linear model controlling for subject age and medication use. Samples (subjects) were ordered to match the metabolite-based ordering from Supplementary Fig. 1. Rows (species) were clustered by Spearman correlation, then grouped by trend with IBD. Units are “relative abundance” (fraction out of 1.0).



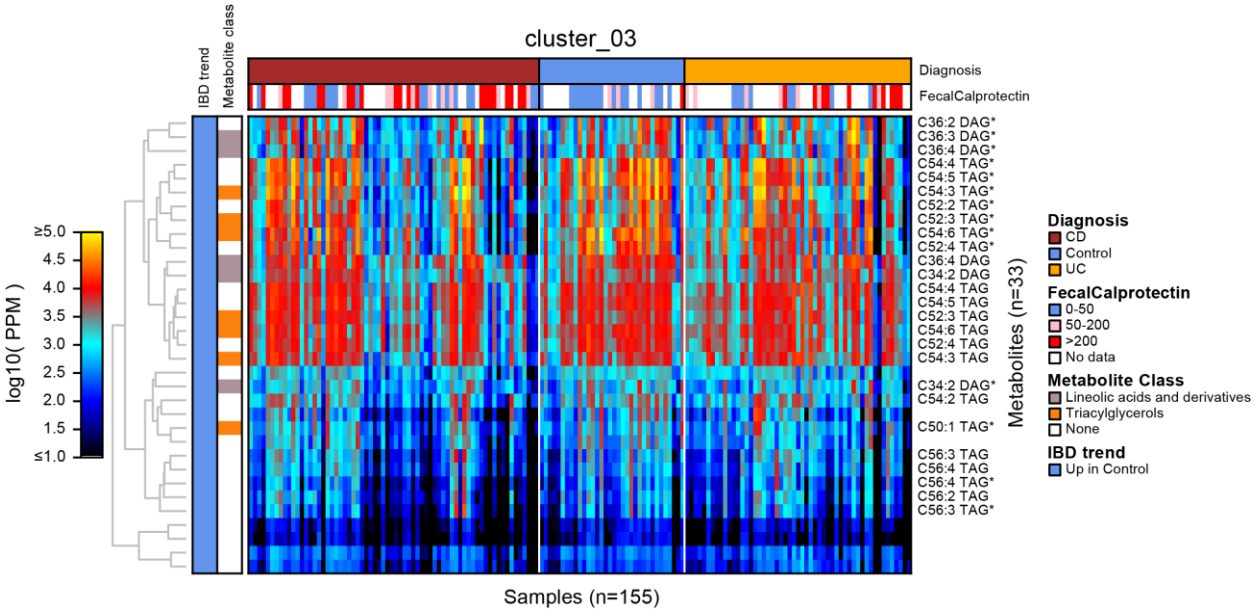
Supplementary Figure 3: IBD is associated with decreases in metabolomic and taxonomic diversity. Panels **A** and **B** replicate Fig. 1, panels B and C from the main text. Panels **C** and **D** compare the first axes of ordination in A and B with subjects' metabolomic and taxonomic Shannon diversity scores, respectively. The metabolomic diversity correlation in C is weak but statistically significant (Spearman's r=-0.321, two-tailed $p<10^{-4}$, n=155), while the taxonomic diversity correlation in D is considerably stronger (r=-0.572, $p<10^{-14}$, n=155).



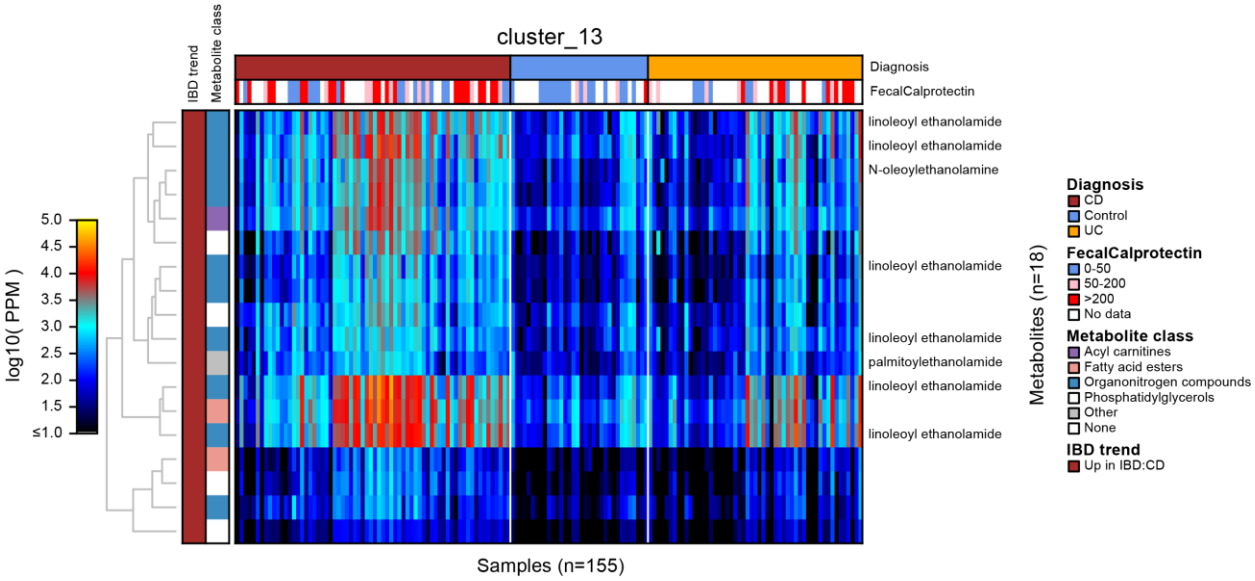
Supplementary Figure 4: Disease localization induces minimal structure on CD subjects' multi'omic profiles. This figure is an analog of Fig. 1, panels B and C from the main text. Samples are ordinated in the same manner, but now colored according to disease localization among CD patients. We confirmed the lack of a significant influence of localization on between-subject distances using permutational analysis of variance (*n*=68, Bray-Curtis distance; *p*=0.22 for metabolites; *p*=0.35 for species).



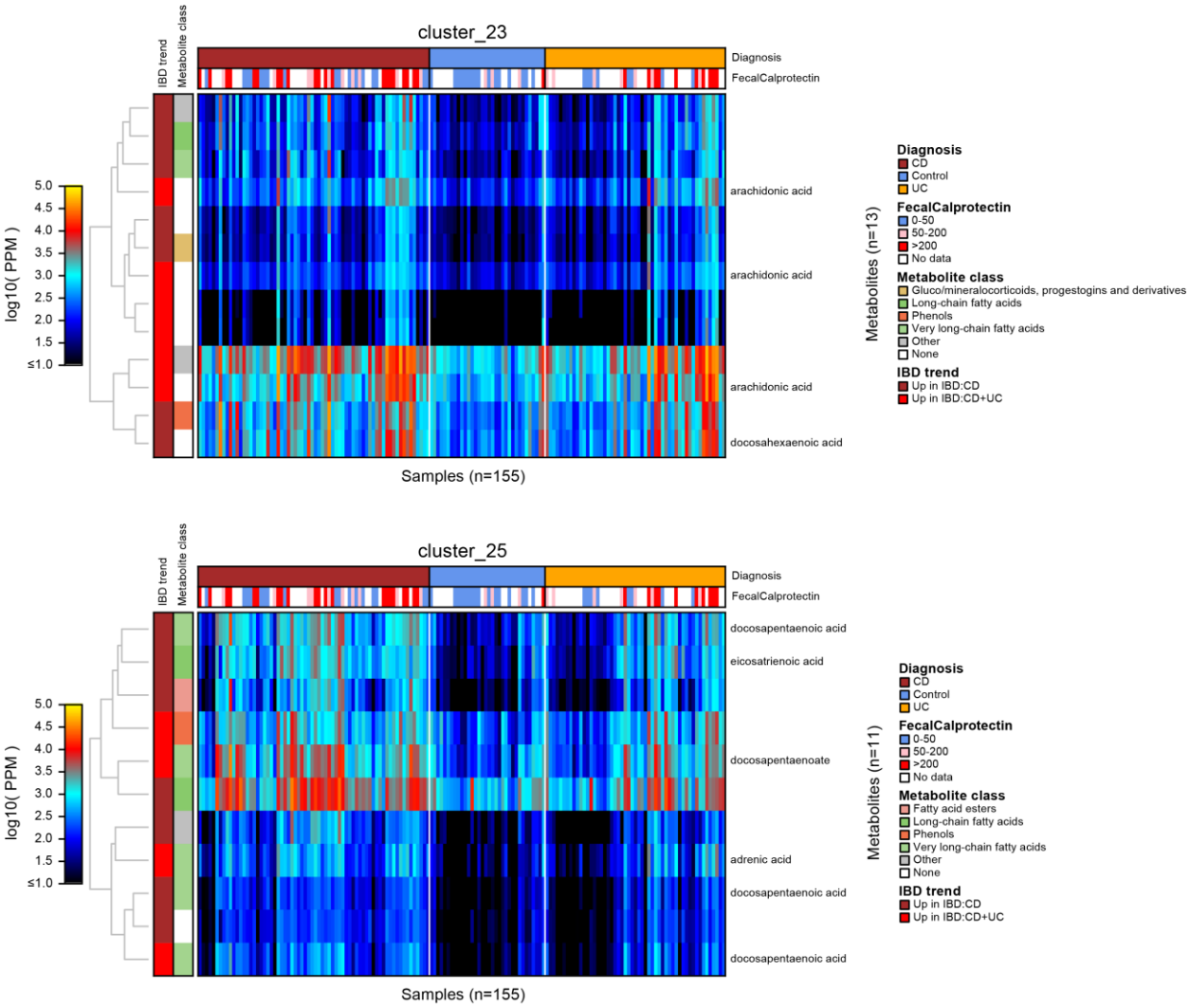
Supplementary Figure 5: Differentially abundant metabolite cluster #3. Differentially abundant metabolite features were clustered (by Spearman correlation) after regressing out the effects of disease, age, and medication use. Clusters were selected with a target intra-cluster similarity *r*=0.7. Clusters tended to be homogeneous with respect to IBD trend, and were often enriched for metabolite features of particular functional classes. Samples (subjects) were ordered to match the metabolite-based ordering from Supplementary Fig. 1. Rows (metabolite features) were clustered by Spearman correlation, then grouped by trend with IBD. Units are "parts per million (PPM)" derived from sum-normalizing metabolomic features intensities within each LC-MS method. Metabolite features are labeled if they were precisely matched against a standard (a "*" indicates a match to one of a group of isomers that could not be differentiated). Cluster #3 was enriched in health, and contained a large number of triacylglycerols.



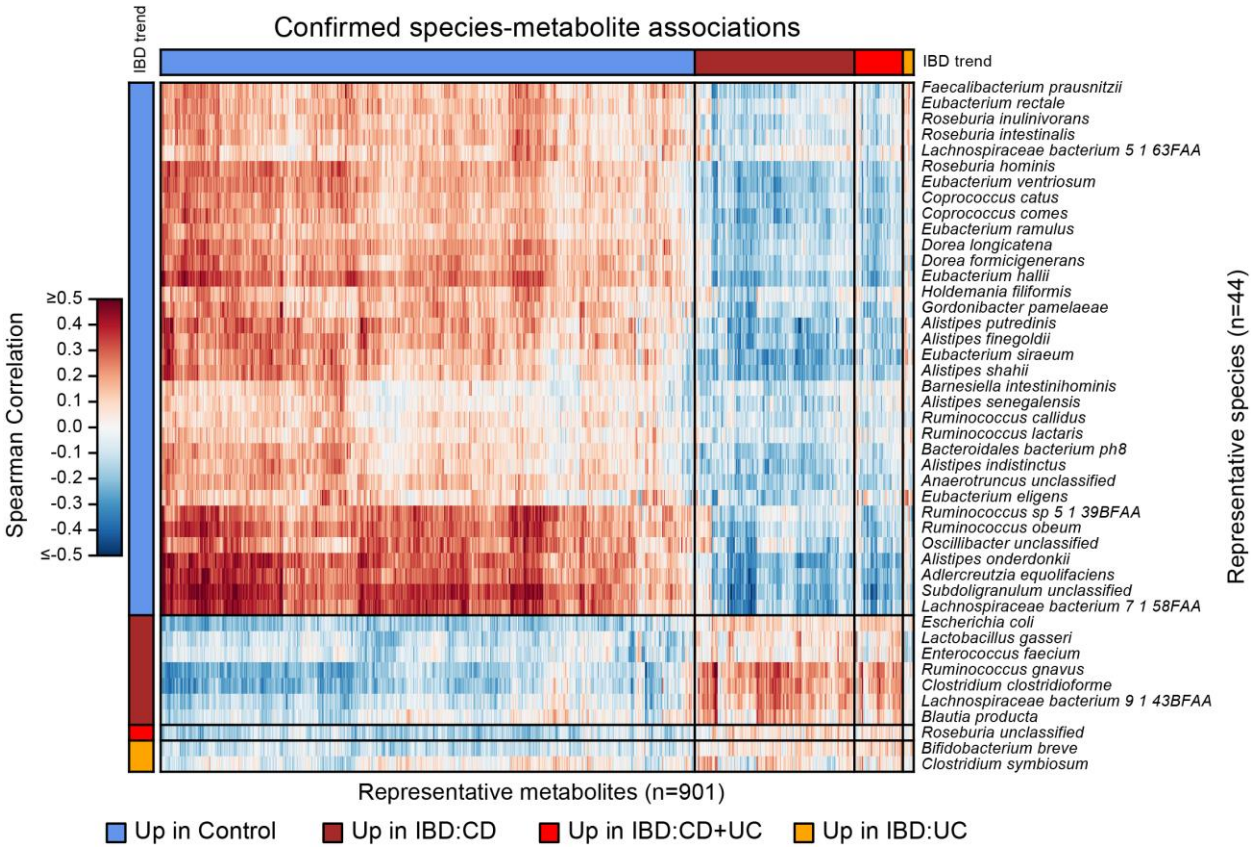
Supplementary Figure 6: Differentially abundant metabolite cluster #13. This figure follows the format of Supplementary Fig. 5 above. Cluster #13 was enriched in CD, and contained a large number of organonitrogen compounds, including metabolite features matched against the standard linoleoyl ethanolamide and its chemical derivatives.



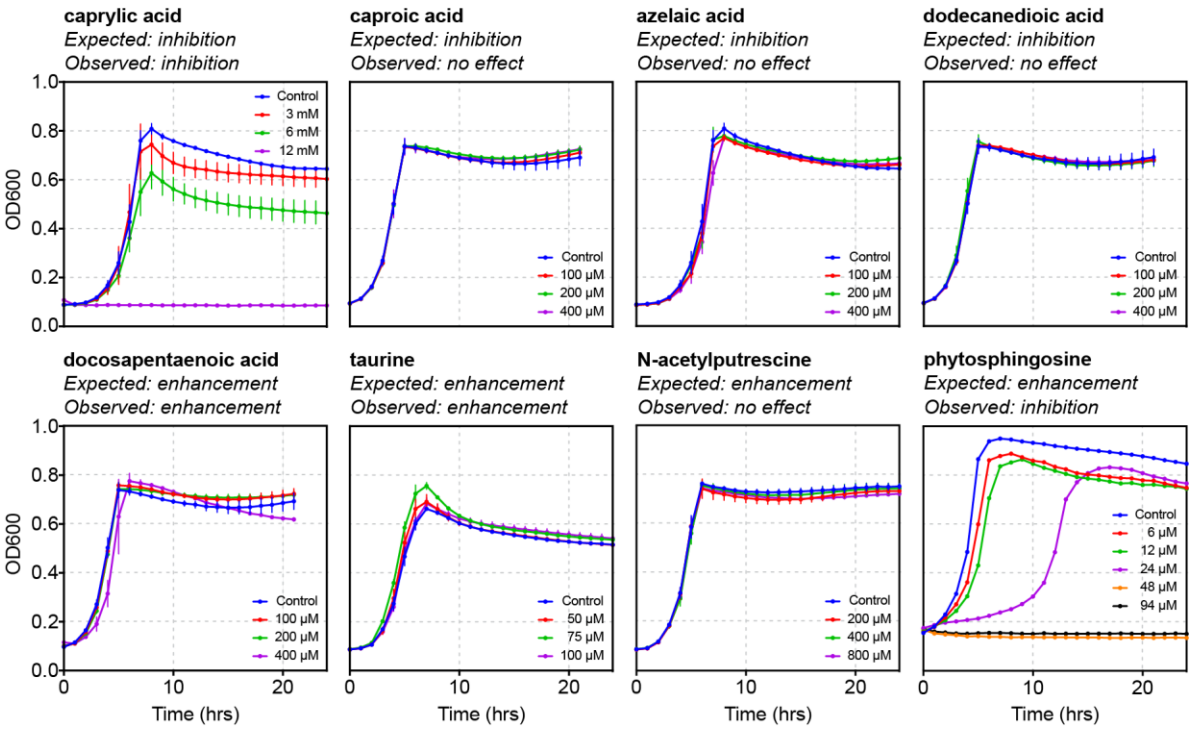
Supplementary Figure 7: Differentially abundant metabolite clusters #23 and #25. This figure follows the format of Supplementary Fig. 5 above. Clusters #23 (top) and #25 (bottom) were enriched in CD and UC, and contained long-chain fatty acids, including metabolite features matched against the standards arachidonic acid and docosapentaenoic acid and their derivatives.



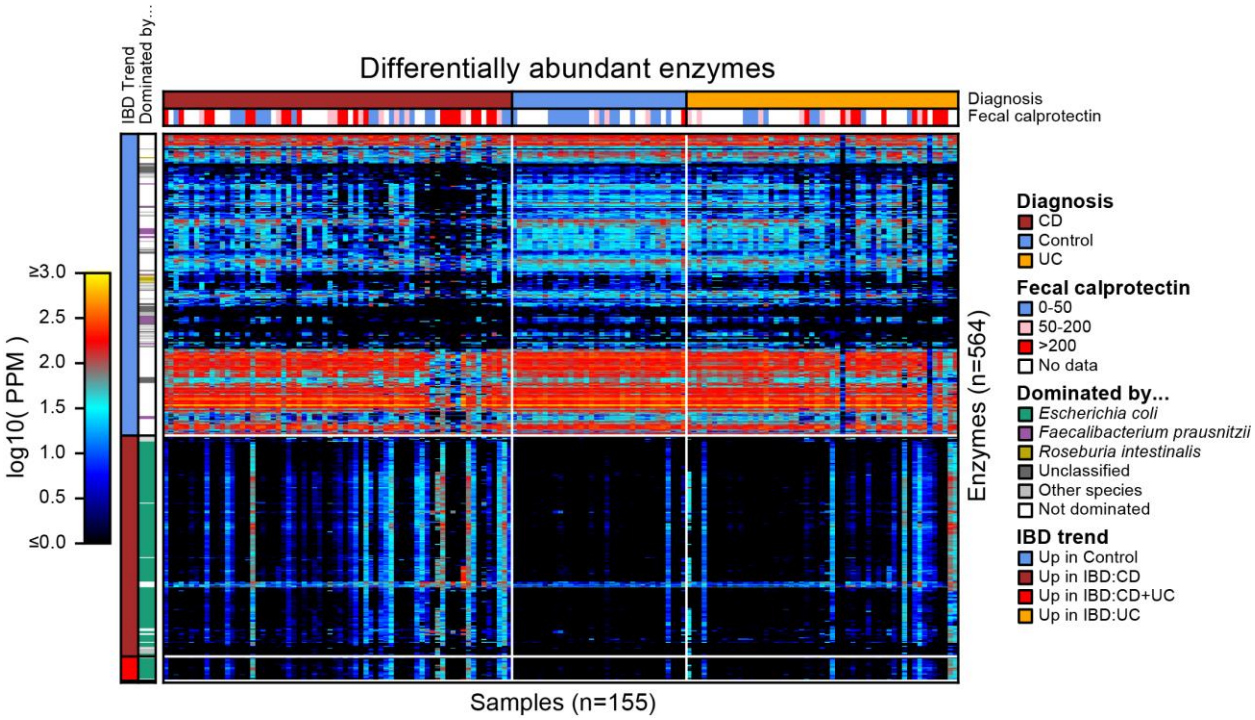
Supplementary Figure 8: Putative mechanistic associations between differentially abundant metabolites and species (expanded view). We identified correlations between species and metabolite features that were individually differentially abundant in IBD. We performed Spearman correlation on abundance data after regressing out the effects of disease, subject age, and medication use. This initial set of correlations was subjected to correction for multiple hypothesis testing (specifically, FDR correction of nominal two-tailed *p*-values with target $q<0.05$). In addition, FDR-significant correlations were filtered to remove associations that were not also 1) nominally significant and 2) of the same sign when considering only raw values from control subjects. This procedure enriches for species-metabolite associations that occur in healthy individuals and which are potentially perturbed in IBD. The figure illustrates the space of correlations between species and metabolite features that were involved in at least one of these putatively mechanistic associations. Note that the majority of associations are concordant with IBD trend (e.g. species and metabolites that are both enriched in IBD tend to correlate positively).



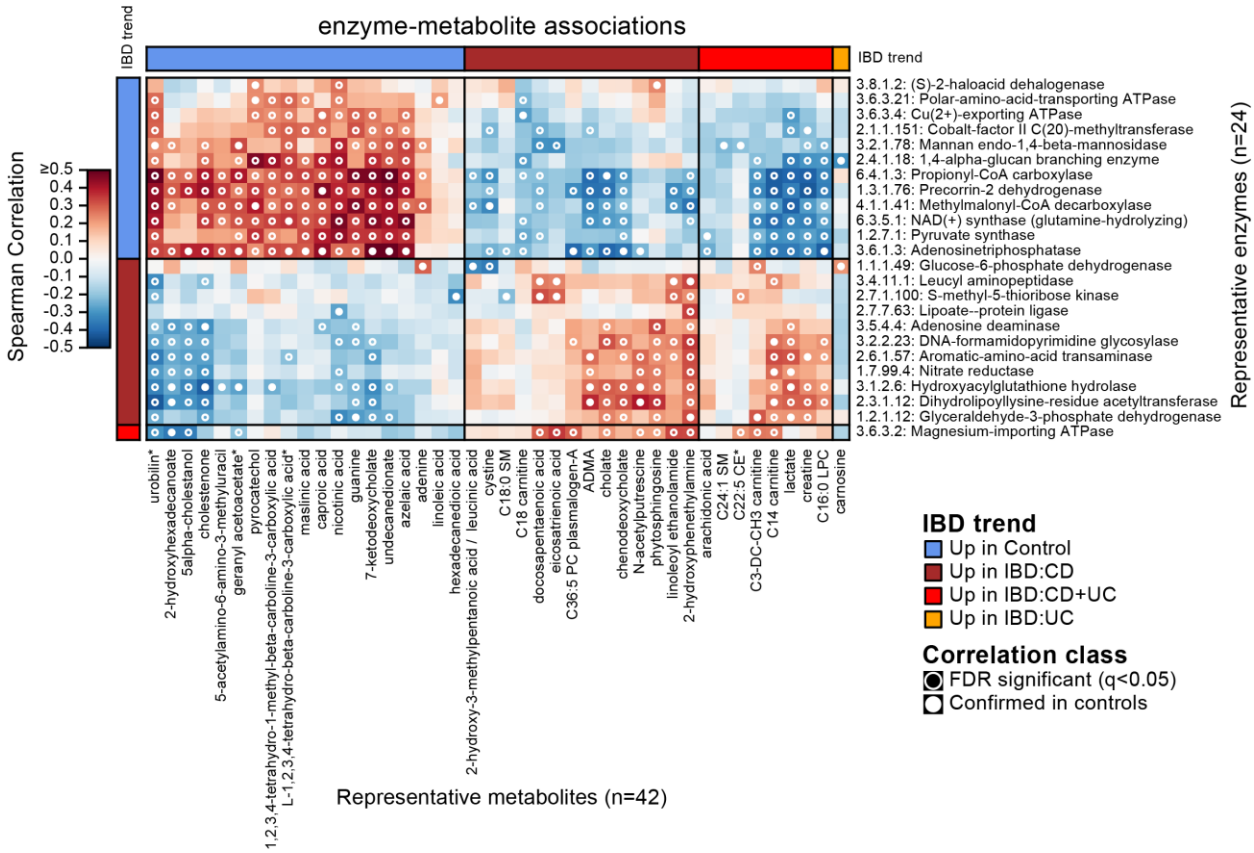
Supplementary Figure 9. Validation of eight predicted metabolite-microbe relationships. We selected eight putative associations between IBD-linked metabolites and the IBD-enriched species *Ruminococcus gnavus* for experimental validation (four negative and four positive). Negatively associated metabolites are predicted to inhibit growth, while positively associated metabolites are predicted to enhance growth. We grew *Ruminococcus gnavus* ATCC 29149 anaerobically in the presence of different concentrations of the indicated metabolites or DMSO control and monitored growth (optical density, OD, at 600 nm) over a 24-hour period. Growth curves representative of two independent tests are shown. Error bars in controls and treated wells represent average ± standard deviation of six and three technical replicates, respectively. Three of the eight observed relationships matched expectations.



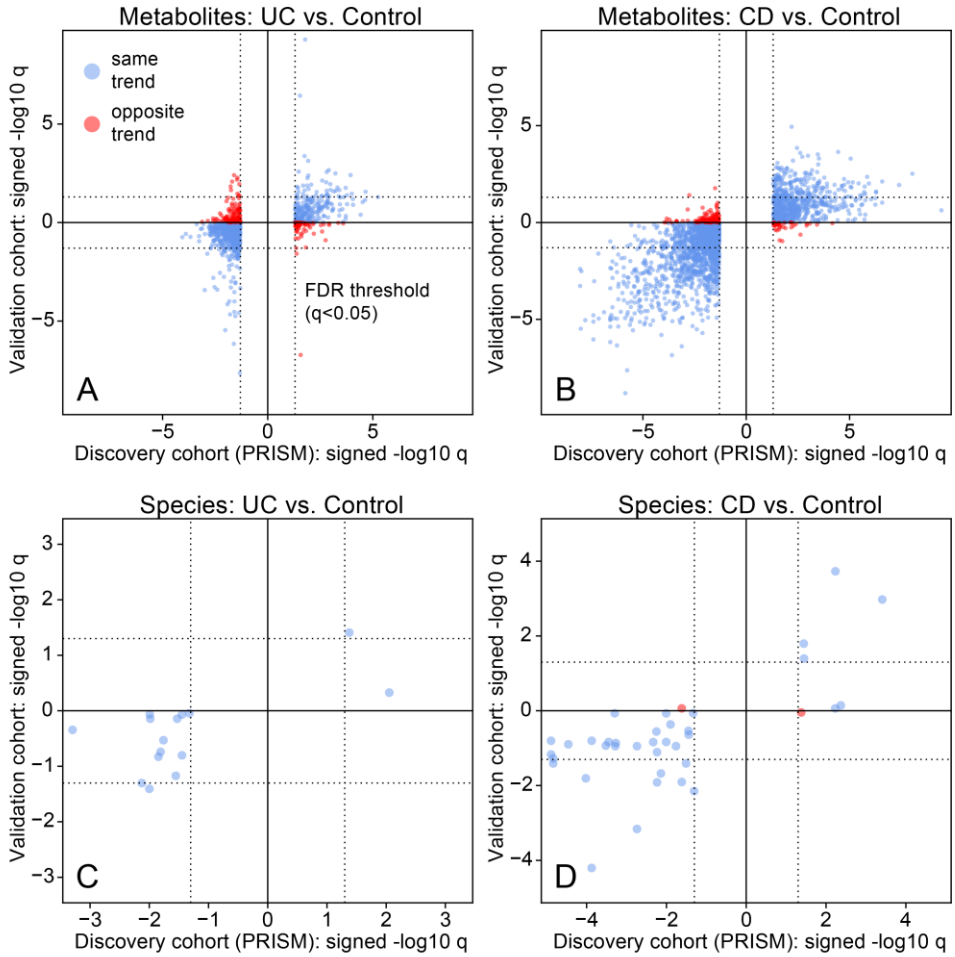
Supplementary Figure 10: Differentially abundant enzymes in IBD. 564 enzymes were differentially abundant IBD (UC and/or CD) relative to controls as determined by a linear model controlling for subject age and medication use. Here, "enzyme" refers to a level-4 category from the Enzyme Commission (EC) hierarchy. Enzymes were quantified by regrouping the abundance of individual protein sequences according to their EC annotations. Samples (subjects) were ordered to match the metabolite-based ordering from Supplementary Fig. 1. Rows (enzymes) were clustered by Spearman correlation, then grouped by trend with IBD. Units are "parts (enzyme copies) per million (PPM)"; here, these units account for the mass of metagenomic reads that did map to any protein sequence, as well as protein sequences lacking EC annotations. Enzymes are colored according to their dominant contributed species, where applicable, with "dominating" defined as "contributing >50% of copies in >50% of samples." Many of the enzymes that are elevated in IBD can be attributed largely to increased levels of *E. coli* in that population.



Supplementary Figure 11: Putative mechanistic associations between differentially abundant metabolites and enzymes. This figure is an analog of Fig. 4 from the main text (which compared species and metabolites). We identified FDR-significant (q<0.05) associations between clustered DA metabolite features and DA enzymes after regressing out variation due to phenotype, age, and medication use (open dots). We selected associations that were also nominally significant among raw control values (filled dots). Here we show the space of metabolite-enzyme correlations involving metabolites matched against standards and characterized species. Spearman's r was used to measure all correlations and all nominal p-values were two-tailed.



Supplementary Figure 12: The majority of IBD-associated trends replicate (in sign) in an independent validation cohort. Each panel compares a set of linear modeling results for a particular test (UC vs non-IBD control or CD vs. non-IBD control) across a particular feature type (metabolite features or microbial species). Plotted values represent the signed, $\log_{10}$-scaled *q*-values of the coefficient following FDR correction. Hence, a large positive value represents a highly significant, positive coefficient (corresponding to a feature that was consistently enriched in IBD relative to controls). Horizontal (*x*) values reflect results from the discovery (PRISM) cohort (*n*=155), and vertical (*y*) values reflect results from the validation (Netherlands) cohort (*n*=65). Dotted lines represent the threshold for FDR-significance (*q*=0.05). Comparisons that were not significant in the discovery cohort are not plotted (hence there are no points between the vertical dotted lines). The vast majority of trends replicated in sign (positive or negative) between the two cohorts (blue points). FDR significance was replicated less often due in part to reduced statistical power in the validation cohort.



Supplementary Figure 13: Classification of IBD status is considerably better than random. This figure expands on Fig. 6A from the main text. Specifically, each training/testing experiment was repeated on 50 permutations of the sample labels. ROC curves from permutation trials are shown in light gray. As expected, the permutation AUC values are centered on 0.5 (random performance), and the true ROC curves (red, blue, and violet lines) fall well outside the permutation "band" (consistent with strong, non-random performance).